

# Structural requirements and discrimination of cell differentiation networks

Christian Breindl<sup>\*,1</sup> Daniella Schittler<sup>\*,1</sup> Steffen Waldherr<sup>\*</sup>  
Frank Allgöwer<sup>\*</sup>

<sup>\*</sup> *Institute for Systems Theory and Automatic Control, University of  
Stuttgart, 70550 Stuttgart, Germany*  
{breindl, schittler, waldherr, allgower}@ist.uni-stuttgart.de

---

**Abstract:** Mathematical models of stem cell differentiation are commonly based upon the concept of subsequent cell fate decisions, each controlled by a gene regulatory network. These networks exhibit a multistable behavior and cause the system to switch between qualitatively distinct stable steady states. However, the network structure of such a switching module is often uncertain, and there is lack of knowledge about the exact reaction kinetics. In this paper, we therefore perform an elementary study of small networks consisting of three interacting transcriptional regulators responsible for cell differentiation: We investigate which network structures can reproduce a certain multistable behavior, and how robustly this behavior is realized by each network. In order to approach these questions, we use a modeling framework which only uses qualitative information about the network, yet allows model discrimination as well as to evaluate the robustness of the desired multistability properties. We reveal structural network properties which are necessary and sufficient to realize distinct steady state patterns required for cell differentiation. Our results also show that structural and robustness properties of the networks are related to each other.

*Keywords:* gene regulatory network, multistability, steady states, feedback loops, switches

---

## 1. INTRODUCTION

Stem cells and their potential to give rise to multiple cell types have become a focus in systems biology and mathematical modeling throughout the last years (Peltier and Schaffer, 2010; MacArthur et al., 2009). Starting from a multi-potent stem cell state, they undergo the process of cell differentiation which is completed when the cell ends up in a mature cell state. The eventual goal in stem cell research is to guide the differentiation of stem cells toward a desired cell type safely, and to maintain this state robustly. Therefore, a promising yet challenging task is to reveal the processes driving cell differentiation by mathematical modeling.

Cell differentiation is commonly viewed and modeled as a sequence of cell fate decisions (Foster et al., 2009). Thereby, the individual decision steps can be represented by modules, each driven by a small gene regulatory network of “master” transcriptional regulators (TRs) in which feedback motifs play a key role (Xiong and Ferrell, 2003; Huang et al., 2007; Foster et al., 2009). The development of models for these modules however is often severely hampered by the lack of detailed knowledge about molecular processes, shifting the focus toward qualitative models in order to understand the fundamental mechanisms. A variety of publications has contributed to the modeling of (stem) cell differentiation, covering a broad range of diverse modeling approaches, e.g. Roeder and Glauche (2006); Huang et al. (2007); Narula et al. (2010); MacArthur et al. (2008). Usually, specific stem cell sys-

tems such as hematopoietic or mesenchymal stem cells are considered. Yet, similar motifs and system properties can be found among them. For example, all these approaches have in common that they represent distinct cell types by distinct stable steady states, declaring multistability as a universal system property.

One of several aspects within this field concerns the question of stable cell states, and minimalistic models that are capable of reproducing experimentally observed cell states. Although it has been established to view cell differentiation as subsequent decisions driven by multistable switching modules (Foster et al., 2009), the network structure of the single modules is often still unclear. We therefore perform an elementary study of small networks consisting of three TRs. These networks should be able to exhibit one progenitor state and two competing differentiated cell types. We also compare two different hypotheses about the steady state levels of the TRs in the individual cell types. In order to reveal properties of 3-node networks that provide candidate models for the differentiation process, we address the following questions:

- Given three cell type-specific TRs, which networks of interactions between the TRs can reproduce the required stable steady states?
- Among the networks that fulfill the stable steady state requirements, are there differences in robustness of the required multistability properties with respect to perturbations in the interactions?

As these are very basic questions, this work builds upon a qualitative modeling framework which needs only very few assumptions about the reaction kinetics. In order to

---

<sup>1</sup> Both authors contributed equally to this work.

model the influence of a TR on its target, monotonous activation and inhibition functions are used, but no exact kinetic parameters have to be specified. This framework has been previously used for example in Chaves et al. (2008); Breindl et al. (2010).

The outline of this paper is as follows: In Section 2, we describe the modeling framework, and define several criteria used to classify the investigated networks. In Section 3, we present and discuss the results of the model analysis. Section 4 gives a short summary and concludes with an outlook on future investigations.

## 2. METHODS

The modeling framework used for this work is briefly recalled first. We then specify the stable steady states to be reproduced by each candidate model, and introduce some characteristic measures and a robustness measure.

### 2.1 Modeling framework

The framework is based on ordinary differential equations and makes use of only very few modeling assumptions which shall be explained next. As we are considering networks of TRs, the influence of such a regulator on its target gene has to be described mathematically. It is assumed that this influence can be described by monotonous activation and inhibition functions. The definitions of these functions are given next.

*Definition 1.* An *activation (inhibition) function* is a function  $\nu : [0, \infty) \rightarrow [0, N)$  ( $\mu : [0, \infty) \rightarrow (0, N]$ ) with  $N \in \mathbb{R}_+$  and:

- i)  $\nu(\mu)$  is continuously differentiable,
- ii)  $\nu(0) = 0$  and  $\nu(x) \rightarrow N$  as  $x \rightarrow \infty$   
( $\mu(0) = N$  and  $\mu(x) \rightarrow 0$  as  $x \rightarrow \infty$ ),
- iii)  $\nu(x)$  ( $\mu(x)$ ) is monotonously increasing (decreasing).

Note, that commonly used kinetics such as Michaelis-Menten or Hill kinetics are covered by this definition. Furthermore, it is assumed that each of the TRs undergoes a linear degradation. With this, the dynamics of a network of  $n$  TRs are given by

$$\dot{x}_i = -k_i \cdot x_i + f_i(x), \quad i = 1, \dots, n. \quad (1)$$

In this,  $x \in \mathbb{R}_+^n$  is the vector of concentrations of the TRs and  $f_i(x)$  models the combined influence of all TRs in a network on the TR  $x_i$ . In this study we only allow multiplicative combinations of activation and inhibition functions. As generally the parameters  $k_i$  and the exact shapes of the activation and inhibition functions are not known, (1) only captures the interaction structure.

As mentioned in the introduction, our goal is to study all possible networks consisting of three nodes  $x_i, i \in \{1, 2, 3\}$ . The interaction structure of such a network is represented by an interaction matrix

$$\mathbf{A} = \{a_{ij}\}_{i,j \in \{1,2,3\}}, \quad a_{ij} := \begin{cases} 1 & \iff x_j \text{ activates } x_i \\ -1 & \iff x_j \text{ inhibits } x_i \\ 0 & \iff \text{no interaction.} \end{cases} \quad (2)$$

To give an example, the interaction matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{bmatrix}$$

encodes the network

$$\begin{aligned} \dot{x}_1 &= -k_1 \cdot x_1 + \nu_1(x_1) \cdot \nu_2(x_3), \\ \dot{x}_2 &= -k_2 \cdot x_2 + \nu_3(x_2) \cdot \mu_4(x_3), \\ \dot{x}_3 &= -k_3 \cdot x_3 + \mu_5(x_1), \end{aligned}$$

in which the parameters  $k_i$  and the shapes of the activation and inhibition functions are not specified.

### 2.2 Specifications of stable steady states

We are specifically interested in the stable steady states of these networks as they determine the cell type. However, in this context we will not consider stable steady states as individual points but as forward-invariant sets in the state space to account for biological variability. By doing so, only high and low levels of a TR are distinguished. To be more precise, we assume concentrations  $x_i^{\text{low}}, x_i^{\text{high}}$  and  $x_i^{\text{max}}$  with  $0 \leq x_i^{\text{low}} \leq x_i^{\text{high}} \leq x_i^{\text{max}}$  to be known, and a concentration  $x_i$  in the interval  $\mathcal{I}_{x_i}^{\text{low}} = [0, x_i^{\text{low}}]$  is considered as low. Equivalently, a concentration  $x_i \in \mathcal{I}_{x_i}^{\text{high}} = [x_i^{\text{high}}, x_i^{\text{max}}]$  is considered as high. With this, a stable steady state of (1) is considered as a forward-invariant hyperrectangular set  $\mathbf{x} = \mathcal{I}_{x_1}^{l_1} \times \dots \times \mathcal{I}_{x_n}^{l_n}$  with  $l_i \in \{\text{high}, \text{low}\}$ . For ease of notation, the  $n$ -tuple of labels  $(l_1, \dots, l_n)$  will be used to refer to this hyperrectangular set.

Let us now turn to our system of three TRs that provide a decision module in the cell differentiation process. In this setup, three steady states are of special interest: The progenitor state  $A$ , a differentiated cell type  $B$ , and a competing differentiated cell type  $C$ . Each of these cell types  $i \in \{A, B, C\}$  is characterized by a specific 3-tuple of high and low TR-levels  $\mathbf{x}^{(i)} = (l_1^{(i)}, l_2^{(i)}, l_3^{(i)})^T$ . One can however find several plausible ways to characterize the three different cell types via levels of the three TRs. We assume the differentiated cell types  $\mathbf{x}^{(B)}, \mathbf{x}^{(C)}$  to be characterized by a high level of the type-specific TR  $x_2$  or  $x_3$ , respectively. Regarding the progenitor state  $\mathbf{x}^{(A)}$ , we aim to compare two hypotheses ("specifications" of stable steady states, in the following for short SSS-specifications) against each other, which represent conceptionally distinct mechanisms:

- (S1)  $x_1$  corresponds to a *progenitor factor*, maintaining the progenitor state.  $x_1$  is high in the progenitor state  $\mathbf{x}^{(A)}$ , whereas it is low in the differentiated states  $\mathbf{x}^{(B)}, \mathbf{x}^{(C)}$ . It has to be *downregulated* in order to achieve cell differentiation. This means the required stable steady states are:

$$\begin{aligned} \mathbf{x}^{(A)} &:= (\text{high}, \text{low}, \text{low}), \\ \mathbf{x}^{(B)} &:= (\text{low}, \text{high}, \text{low}), \\ \mathbf{x}^{(C)} &:= (\text{low}, \text{low}, \text{high}). \end{aligned} \quad (3)$$

- (S2)  $x_1$  corresponds to a *differentiation factor*, enabling differentiation.  $x_1$  is low in the progenitor state  $\mathbf{x}^{(A)}$ , whereas it is high in the differentiated states  $\mathbf{x}^{(B)}, \mathbf{x}^{(C)}$ . It has to be *upregulated* in order to achieve cell differentiation. This corresponds to requiring the stable steady states:

$$\begin{aligned}\mathbf{x}^{(A)} &:= (\text{low}, \text{low}, \text{low}), \\ \mathbf{x}^{(B)} &:= (\text{high}, \text{high}, \text{low}), \\ \mathbf{x}^{(C)} &:= (\text{high}, \text{low}, \text{high}).\end{aligned}\quad (4)$$

Besides the specified stable steady states, we impose a further requirement on the candidate models: In order to investigate “truly interacting” nodes rather than assemblages of isolated nodes or subgraphs, we require each 3-node network to be weakly connected.

### 2.3 Characteristic measures

Several characteristic measures are introduced to classify networks with respect to their interactions, including self-loops, and the types of these interactions (activating / inhibiting). The *number of nonzero entries* in the interaction matrix  $\mathbf{A}$ , which equals the zero-norm  $\|\mathbf{A}\|_0$ , is used as a measure of the network's connectivity. To measure the preponderance of activating versus inhibiting interactions, we use the number of positive entries in  $\mathbf{A}$  minus the number of negative entries in  $\mathbf{A}$ , which we call the *positive-negative weight*.

The *number of sign-inconsistent (sign-consistent) loops* is used to measure the degree of inconsistencies (consistencies, respectively) arising from network interactions. A loop is either a feedback loop from  $x_i$  to itself, or a feedforward loop from  $x_i$  to  $x_j$ . Sign-inconsistencies arise when a feedback loop has negative sign, or the two paths of a feedforward loop from  $x_i$  to  $x_j$  have opposite signs. Sign-consistent loops are feedback loops with positive sign, as well as feedforward loops with the two paths having the same sign. Similar motifs have been outlined e.g. in Ma'ayan et al. (2008).

To measure how strongly the connectivity is distributed among the three nodes, the *maximum indegree* is computed for each network:

$$\text{maxindeg}(\mathbf{A}) := \max_{i \in \{1,2,3\}} \left\{ \sum_{j \in \{1,2,3\}} a_{ij} \right\}, \quad (5)$$

which is the largest row-sum, thus the  $\infty$ -norm  $\|\mathbf{A}\|_\infty$ .

### 2.4 Robustness measure

Since the goal of this study is to compare networks that can reproduce the desired cell types, we computed for each network structure a robustness measure  $\mathcal{R}$  that quantifies if and how well this structure can generate a set of forward-invariant sets as specified in (3,4). We applied the algorithm introduced in Breindl et al. (2010) which is summarized next.

In order to explain the underlying idea of this method, let us use the symbol  $\varphi$  to denote both, activation and inhibition functions. These functions are indexed  $\varphi_{i,k}$  such that  $i$  denotes the regulated TR, i.e.,  $x_i$ , and  $k$  enumerates the regulators of  $x_i$ . The number of regulators of  $x_i$  is denoted as  $q_i$ .

Also, let us measure a perturbation of a monotonous function  $\varphi$  as the  $l_1$ -norm of the difference of the original monotonous function  $\varphi$  and the perturbed monotonous function  $\varphi^p$ , i.e.,

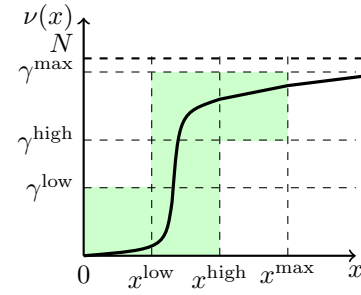


Fig. 1. Illustration of a tube for an activation function.

$$\|\varphi - \varphi^p\|_1 = \int_0^\infty |\varphi(x) - \varphi^p(x)| dx. \quad (6)$$

Given a set of desired forward-invariant sets  $\mathbf{x}^z$ ,  $z = 1, \dots, m$ , the goal of the method is to assign a robustness value  $\mathcal{R}$  to the interaction structure itself, i.e., to the matrix  $\mathbf{A}$  as in (2). This robustness value  $\mathcal{R}$  is defined such that it has the following three properties: (i) If a value  $\mathcal{R}$  is assigned to a system  $\mathbf{A}$ , there exists a realization of  $\mathbf{A}$  with monotonous functions  $\tilde{\varphi}_{i,k}$ , such that the desired sets  $\mathbf{x}^z$ ,  $z = 1, \dots, m$ , are forward-invariant. (ii) If no monotonous function is perturbed by more than  $\mathcal{R}$ , i.e.,  $\forall i, k : \|\tilde{\varphi}_{i,k} - \varphi_{i,k}^p\|_1 \leq \mathcal{R}$ , it can be guaranteed that forward-invariance of the sets  $\mathbf{x}^z$ ,  $z = 1, \dots, m$ , is maintained. (iii) For every  $\hat{\mathcal{R}} > \mathcal{R}$ , there exist perturbed functions  $\hat{\varphi}_{i,k}^p$  such that forward-invariance of the sets  $\mathbf{x}^z$ ,  $z = 1, \dots, m$ , is lost, and for at least one index  $i, k$  it holds that  $\|\tilde{\varphi}_{i,k} - \hat{\varphi}_{i,k}^p\|_1 = \hat{\mathcal{R}}$ .

The procedure to compute  $\mathcal{R}$  for a given network structure  $\mathbf{A}$  and specification  $\mathbf{x}^z$ ,  $z = 1, \dots, m$ , is outlined next. To this end, two definitions are given next.

**Definition 2.** The 3-tuple of pairs of positive real numbers  $T_\nu = ((x^{\text{low}}, \gamma^{\text{low}}), (x^{\text{high}}, \gamma^{\text{high}}), (x^{\text{max}}, \gamma^{\text{max}}))$  such that  $\gamma^{\text{low}} \leq \gamma^{\text{high}} \leq \gamma^{\text{max}}$  and  $x^{\text{low}} \leq x^{\text{high}} \leq x^{\text{max}}$  is called tube for an activation function. Equivalently, the 3-tuple of pairs of positive real numbers  $T_\mu = ((x^{\text{low}}, \gamma^{\text{high}}), (x^{\text{high}}, \gamma^{\text{low}}), (x^{\text{max}}, \gamma^{\text{min}}))$  such that  $\gamma^{\text{min}} \leq \gamma^{\text{low}} \leq \gamma^{\text{high}}$  and  $x^{\text{low}} \leq x^{\text{high}} \leq x^{\text{max}}$  is called tube for an inhibition function.

**Definition 3.** An activation function  $\nu$  (inhibition function  $\mu$ ) is said to satisfy a tube  $T_\nu$  ( $T_\mu$ ), denoted as  $\nu \models T_\nu$  ( $\mu \models T_\mu$ ) if the following inequalities hold.

$$\forall x \leq x^{\text{low}} : \nu(x) \leq \gamma^{\text{low}} \quad (\mu(x) \geq \gamma^{\text{high}}) \quad (7)$$

$$\forall x \geq x^{\text{high}} : \nu(x) \geq \gamma^{\text{high}} \quad (\mu(x) \leq \gamma^{\text{low}}) \quad (8)$$

$$\forall x \leq x^{\text{max}} : \nu(x) \leq \gamma^{\text{max}} \quad (\mu(x) \geq \gamma^{\text{min}}) \quad (9)$$

This means that a tube for a monotonous function restricts the space where the graph of this function can evolve (see also Figure 1). Note that the  $x$ -values of the tubes are given by the specification of the forward-invariant sets. It is now possible to formulate conditions on the  $\gamma$ -values of the tubes such that the specified sets are forward-invariant if all functions lie inside their tubes, i.e.  $\forall i, k : \varphi_{i,k} \models T^{i,k}$ , where  $T$  can mean  $T_\nu$  or  $T_\mu$  as required, and the tubes are indexed as the according monotonous functions.

**Proposition 1.** (Breindl et al. (2010)). Given a set  $\mathbf{x} = \mathcal{I}_{x_1}^{l_1} \times \dots \times \mathcal{I}_{x_n}^{l_n}$  with  $l_i \in \{\text{low}, \text{high}\}$  and given tubes  $T^{i,k}$  that satisfy the conditions

$$\forall i \in \{1, \dots, n\} : -k_i \cdot x_i + \gamma_{i,1} \cdot \dots \cdot \gamma_{i,q_i} \geq 0 \quad (10)$$

where  $\underline{x}_i = \min_{x_i \in \mathcal{I}_{x_i}^{l_i}} x_i$ , and, with  $x_j$  denoting the argument of  $\varphi_{i,k}$ ,

$$\underline{\gamma}_{i,k} = \begin{cases} 0 & \text{if } \varphi_{i,k} = \nu_{i,k} \wedge 0 \in \mathcal{I}_{x_j}^{l_j} \\ \min\{\gamma : (x, \gamma) \in T^{i,k} \wedge x \in \mathcal{I}_{x_j}^{l_j}\} & \text{otherwise} \end{cases}$$

and

$$\forall i \in \{1, \dots, n\} : -k_i \cdot \bar{x}_i + \bar{\gamma}_{i,1} \cdot \dots \cdot \bar{\gamma}_{i,q_i} \leq 0 \quad (11)$$

where  $\bar{x}_i = \max_{x_i \in \mathcal{I}_{x_i}^{l_i}} x_i$ , and, with  $x_j$  denoting the argument of  $\varphi_{i,k}$ ,

$$\bar{\gamma}_{i,k} = \max\{\gamma : (x, \gamma) \in T^{i,k} \wedge x_j \in \mathcal{I}_{x_j}^{l_j}\}.$$

If  $\forall i, k : \varphi_{i,k} \models T^{i,k}$ , then the set  $\mathbf{x}$  is forward-invariant for system (1).

Starting from this result, the robustness measure can now be introduced. As forward-invariance for the specified sets can be guaranteed as long as  $\forall i, k : \varphi_{i,k} \models T^{i,k}$ , the goal is to find, for each tube  $T^{i,k}$ , the function  $\tilde{\varphi}_{i,k}$ , which is best centered to the tube  $T^{i,k}$ . In this context, best centered means that  $\tilde{\varphi}_{i,k}$  has to be perturbed more than any other function  $\varphi_{i,k} \models T^{i,k}$  in order to violate at least one of the constraints (7-9) (denoted as  $\varphi_{i,k} \not\models T^{i,k}$ ). The minimal perturbation of this best centered function  $\tilde{\varphi}_{i,k}$  in order to achieve a violation of the tube constraints is given as the result of the optimization problem

$$\mathcal{R}^{\max}(T^{i,k}) = \sup_{\varphi_{i,k} \models T^{i,k}} \mathcal{R}^{\min}(\varphi_{i,k}, T^{i,k}), \quad (12)$$

with

$$\mathcal{R}^{\min}(\varphi_{i,k}, T^{i,k}) = \inf_{\varphi_{i,k}^p \not\models T^{i,k}} \|\varphi_{i,k} - \varphi_{i,k}^p\|. \quad (13)$$

Then, the robustness measure is obtained by maximizing the smallest value  $\mathcal{R}^{\max}(T^{i,k})$  of all tubes in the network over all tubes  $T^{i,k}$  that satisfy Proposition 1.

**Definition 4.** Given a system (1) with unspecified activation and inhibition functions. The robustness measure  $\mathcal{R}$  for the system is defined as

$$\begin{aligned} \mathcal{R} &= \max_{T^{i,k}} \min_{i,k} \mathcal{R}^{\max}(T^{i,k}) \\ \text{s.t.: } \forall \mathbf{x}^z : (10) \text{ and } (11) \text{ hold.} \end{aligned} \quad (14)$$

This measure has the interpretation given at the beginning of this section. Furthermore, if there exists no realization of  $\mathbf{A}$  such that the desired sets are forward-invariant, the optimization problem (14) is infeasible. It could furthermore be shown that for networks where the functions  $f_i(x)$  are products of activation and inhibition functions, problem (14) results in a convex optimization problem (Breindl et al., 2010).

### 3. RESULTS

#### 3.1 General observations

Constructing all possible 3-node networks with interactions as specified in (2), i.e. each interaction having exactly one value  $\in \{-1, 0, 1\}$ , yields  $3^{3 \cdot 3} = 19683$  possible networks. Requiring the networks to be weakly connected leaves 19008 networks. Out of these candidate networks, the algorithm from Section 2.4 identifies 206 models that meet the specification (S1), i.e. 1.08% out of all networks; and 242 models that are able to reproduce the specification (S2), i.e. 1.27% out of all networks.

The distributions of the characteristic measures as outlined in Section 2 are shown in Fig. 2. Models for (S1) (Fig. 2 upper row, red) and models for (S2) (Fig. 2 upper row, green) show similar distributions of nonzero entries. The distribution of the positive-negative weight for models of (S1) is shifted to the left, i.e. toward more negative entries, whereas for models of (S2) it is slightly shifted to the right, i.e. toward more positive entries. These differences are in accordance with the number of inconsistent loops which require negative entries: Models for (S1) show a distribution of inconsistent loops up to high numbers and only intermediate amount of consistent loops, whereas models for (S2) tend to low numbers of inconsistencies and more consistent loops. Also, models for (S1) (red) are found for all numbers of inconsistent loops, whereas models for (S2) (green) are restricted to up to 8 inconsistent loops. In contrast, the (S1)-models are more restricted to certain numbers of consistent loops than the (S2)-models are.

Summarizing, there are less networks supporting the hypothesis of a progenitor factor (S1), and they show a tendency toward negative (inhibiting) interactions and hence inconsistent loops. The number of networks providing models for the hypothesis of a differentiation factor (S2) is higher. On average, these networks have about equal number of positive and negative entries, thus reducing the amount of inconsistent loops in favor of consistent loops.

#### 3.2 Structural properties

Regarding the network structure, we found that all models that can reproduce a SSS-specification can be classified leading to sufficient and necessary conditions on the interaction structure. These findings are summarized in the following observations.

**Observation 1.** Let  $\mathbf{A}$  be an interaction matrix (2). If the network represented by  $\mathbf{A}$  is such that each node  $x_i$ ,  $i \in \{1, 2, 3\}$  together with its incoming links is contained in Fig. 3, then there exists a realization of  $\mathbf{A}$  that can reproduce the SSS-specification (S1). Also the other direction holds: For each weakly connected 3-node network that can be assembled from nodes  $x_i$  together with its incoming links contained in Fig. 3 there exists a realization that can reproduce the SSS-specification (S1).

An equivalent observation can be made for the network that can reproduce the SSS-specification (S2).

**Observation 2.** Let  $\mathbf{A}$  be an interaction matrix (2). If the network represented by  $\mathbf{A}$  is such that the node  $x_1$  and the nodes  $x_j$ ,  $j \in \{2, 3\}$  together with their incoming links are contained in Fig. 3, then there exists a realization of  $\mathbf{A}$  that can reproduce the SSS-specification (S1). Also the other direction holds: For each weakly connected 3-node network that can be assembled from nodes  $x_1$  and  $x_j$ ,  $j \in \{2, 3\}$ , together with their incoming links contained in Fig. 3 there exists a realization that can reproduce the SSS-specification (S2).

From Fig. 3 it can also be explained why the positive-negative weight of models for SSS-specification (S1) tends toward a larger number of negative entries: All interactions between two different nodes in the network need to be negative (inhibiting). Only self-loops are allowed to be activating. For SSS-specification (S2) on the other side also activating links from  $x_1$  to the other nodes are possible.

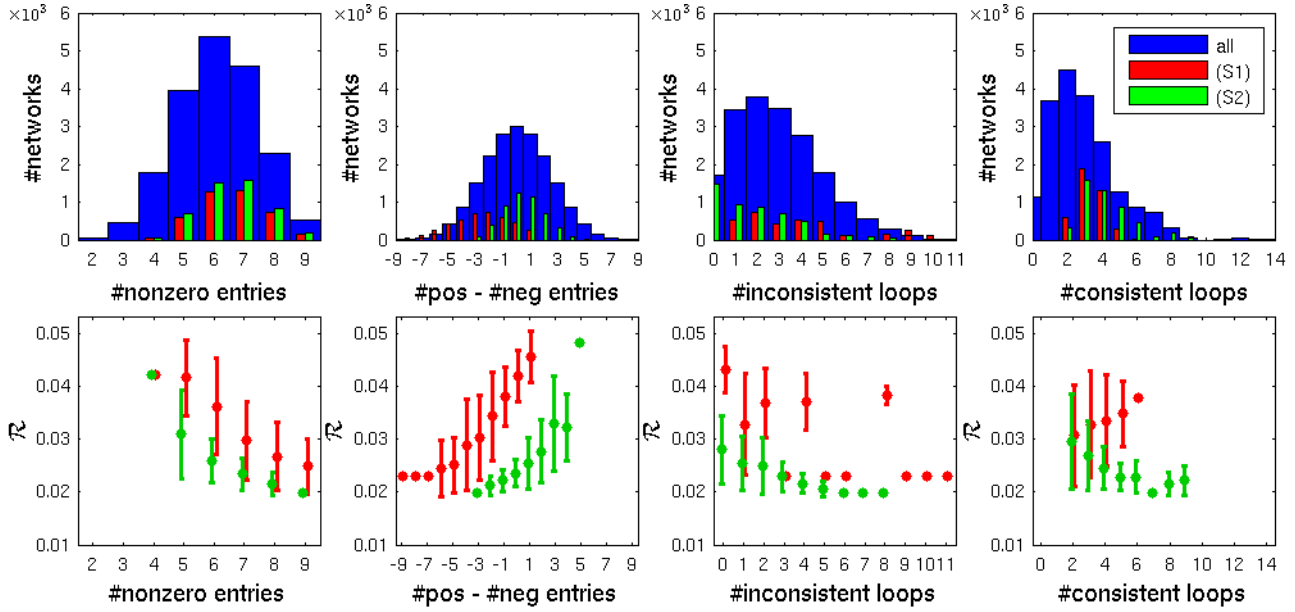
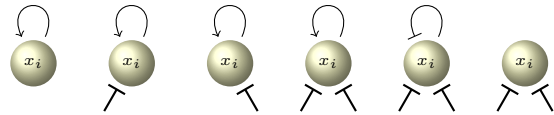
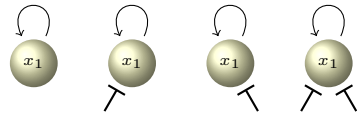


Fig. 2. (Upper row) Distributions of characteristic measures of all 19008 possible weakly-connected 3-node networks (blue), of the 206 models reproducing the SSS-specification (S1) (red,  $\times 20$ ), and of the 242 models reproducing (S2) (green,  $\times 20$ ). (Lower row) Mean robustness values for each characteristic measure, over the 206 models reproducing (S1) (red), and over the 242 models reproducing (S2) (green).

(S1)-models, every node  $x_i$ ,  $i \in \{1, 2, 3\}$



(S2)-models, node  $x_1$



(S2)-models, nodes  $x_j$ ,  $j \in \{2, 3\}$

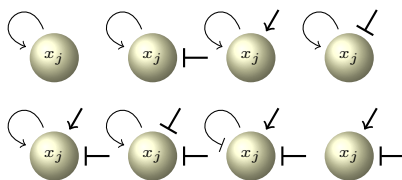


Fig. 3. Complete enumeration of all “building blocks”. Networks assembled from these building blocks are able to reproduce the respective SSS-specification. (Upper row:) Links entering  $x_i$  from bottom left and right represent links from nodes  $x_j$ ,  $j \neq i$  and  $x_k$ ,  $i \neq k \neq j$ . (Middle row:) Links entering  $x_1$  from below left and right represent links from nodes  $x_2$  and  $x_3$ . (Lower row:) Links entering from top right represent links from  $x_1$ . Links entering from right represent links from  $x_k$ ,  $k \in \{2, 3\}$ ,  $k \neq j$ .

Using these “building blocks” as listed in Fig. 3 it is possible to construct all possible 3-node models that can, for an appropriate choice of activation and inhibition functions, reproduce the SSS-specification (S1) or (S2). To see possible applications of our approach, note that a switching module for mesenchymal stem cell differentiation based on literature data and analyzed in detail in

Schittler et al. (2010) satisfies specification (S1) and can indeed be composed from the building blocks presented in Fig. 3. By applying the algorithm from Section 2.4, it is in principle possible to compute the building blocks for any SSS-specification. The algorithm is also directly applicable to larger networks.

### 3.3 Robustness properties

The mean robustness value  $\mathcal{R}$  for all characteristic measures is shown in Fig. 2 (lower row). Two important observations can be made: There is a clear negative correlation of  $\mathcal{R}$  with increasing nonzero entries, and a clear positive correlation of  $\mathcal{R}$  with increasing positive-negative weight. Both observations are true independent of the specific SSS-specifications (red, green). In contrast, no clear correlation of  $\mathcal{R}$  versus the number of inconsistent or consistent loops can be detected.

It is stated by Kwon et al. (2008) that perturbations on nodes involved in fewer (more) than the average number of feedback loops have a lower (higher, respectively) impact on the overall network robustness. Thus, we investigated whether there is a similar relationship between the *maxindeg* and the robustness value  $\mathcal{R}$ . The results of our analysis are shown in Fig. 4. Among all (S1)-reproducing models, only four different robustness values have been assigned while for the (S2)-reproducing models seven different robustness values have been computed. Furthermore, the (S1)-models are more abundant in higher robustness values, whereas the majority of (S2)-models has a lower robustness value (cf. Fig. 4). That is, models for hypothesis (S2) require a finer “tuning” than models for (S1), biologically pointing toward more vulnerable regulatory networks.

Yet, despite these differences, the similarities between the two distributions are eminent which suggests a strong relation between the network’s structural properties and

their robustness: Independent of the specification (S1) or (S2), networks with a higher *maxindeg* tend to have lower  $\mathcal{R}$  values. This means that, the more incoming interactions a node has, the more vulnerable the network gets to perturbations of these interaction. This is in accordance with the findings of Kwon et al. (2008) about network structure and their notion of robustness.

#### 4. CONCLUSION

The main focus of this work was to analyze which networks consisting of three transcriptional regulators are in principle able to reproduce a set of required stable steady states, to search for interaction patterns that are necessary in order to meet these requirements, and to compute and compare the robustness properties of these networks. Still, the investigations raise new questions and possible future work, which we point out briefly in this concluding outlook.

Sign-inconsistencies make a system non-monotone, and the more inconsistent loops are contained in a network, the “farther” the system is from monotonicity. Ma’ayan et al. (2008) argue that intracellular systems will be “close-to-monotone”. Kwon et al. (2008) report that a higher number of positive feedback loops and a smaller number of negative feedback loops result in a higher robustness. These arguments are in accordance with the result here that positive (negative) entries increase (decrease) the robustness of a network. But in contrast to the number of positive / negative entries, the number of inconsistent and consistent loops does not show a clear effect on the robustness as defined here. What is more, inconsistent loops are even required for some SSS-specifications to be met. These studies use different notions of robustness, and their interrelation should be worth further investigations.

Furthermore, one might argue that especially in cell differentiation, stimulus inputs and transitions between stable steady states rather than just the existence of stable steady states play an important role. However, although these aspects were not considered in this study, the approach in this paper provides a first selection step of networks with respect to one necessary condition (out of several), namely by checking whether the required stable steady states can be reproduced. These aspects provide plenty of opportunities for separate investigations, and we hope to address these issues in future work.

#### ACKNOWLEDGEMENTS

The authors would like to thank for financial support from the German Research Foundation (DFG) within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart, and from the German Federal Ministry of Education and Research (BMBF) within the SysTec program (grant nr. 0315506A), and from the Helmholtz Society within the CoReNe project. D.S. acknowledges financial support by The MathWorks Foundation of Science and Engineering.

#### REFERENCES

- C. Breindl, S. Waldherr and F. Allgöwer. A robustness measure for the stationary behavior of qualitative gene regulation networks. Proceedings of the 11th symposium on computer applications in biotechnology (CAB), pp. 36–41, 2010.
- M. Chaves, T. Eissing and F. Allgöwer. Bistable Biological Systems: A Characterization Through Local Compact Input-to-State Stability. IEEE Transactions on Automatic Control, Special Issue on Systems Biology, vol. 53, pp. 87–100, 2008.
- D.V. Foster, J.G. Foster, S. Huang and S.A. Kauffman. A model of sequential branching in hierarchical cell fate determination. Journal of Theoretical Biology, vol. 260, pp. 589–597, 2009.
- S. Huang, Y.-P. Guo, G. May and T. Enver. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. Developmental Biology, vol. 305, pp. 695–713, 2007.
- Y.-K. Kwon and K.-H. Cho. Quantitative analysis of robustness and fragility in biological networks based on feedback dynamics. Bioinformatics, vol. 24, pp. 987–994, 2008.
- A. Ma’ayan, A. Lipshtat, R. Iyengar and E.D. Sontag. Proximity of intracellular regulatory networks to monotone systems. IET Systems Biology, vol. 2, pp. 103–112, 2008.
- B.D. MacArthur, C.P. Please and R.O.C. Oreffo. Stochasticity and the molecular mechanisms of induced pluripotency. PLoS ONE, vol. 3, e3086, 2008.
- B.D. MacArthur, A. Ma’ayan and I.R. Lemischka. Systems biology of stem cell fate and cellular reprogramming. Nature Reviews Molecular Cell Biology, vol. 10, pp. 672–681, 2009.
- J. Narula, A.M. Smith, B. Gottgens and O.A. Igoshin. Modeling reveals bistability and low-pass filtering in the network module determining blood stem cell fate. PLoS Computational Biology, vol. 6, e1000771, 2010.
- J. Peltier and D.V. Schaffer. Systems biology approaches to understand stem cell fate choice. IET Systems Biology, vol. 4, pp. 1–11, 2010.
- I. Roeder and I. Glauche. Towards an understanding of lineage specification in hematopoietic stem cells: A mathematical model for the interaction of transcription factors GATA-1 and PU.1. Journal of Theoretical Biology, vol. 241, pp. 852–865, 2006.
- D. Schittler, J. Hasenauer, F. Allgöwer, and S. Waldherr. Cell differentiation modeled via a coupled two-switch regulatory network. Chaos, vol. 20(4):045121, 2010.
- W. Xiong and J.E. Ferrell, Jr.. A positive-feedback-based bistable memory module that governs a cell fate decision. Nature, vol. 426, pp. 460–465, 2003.

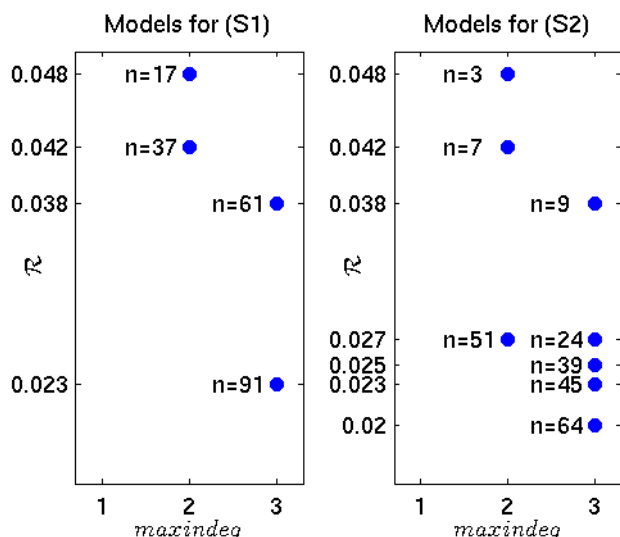


Fig. 4. The robustness value  $\mathcal{R}$  versus the *maxindeg* for all models. For each (*maxindeg*,  $\mathcal{R}$ ), there are  $n$  models which have these values of *maxindeg*,  $\mathcal{R}$ . (Left) For specification (S1), there are four different robustness values, each of them uniquely assigned to a *maxindeg* value. (Right) For specification (S2), there are seven different robustness values. The assignment of a given  $\mathcal{R}$  to a *maxindeg* value is not unique.